CONVERGENCE THEOREMS FOR NONLINEAR APPROXIMATION ALGORITHMS

R. Schaback

Institut für Numerische und Angewandte Mathematik
der Universität Göttingen
Göttingen, FRG

The author gratefully dedicates this paper to the memory of his academic teacher, Prof. Dr. Helmut Werner.

Abstract. *For a large class of algorithms, including trust-region-, projection-, line-search-, and Levenberg-Marquardt methods, a quantitative global convergence proof based on a general convergence theory is given.*

1 Introduction

We use the notations and conventions of M.J.D. POWELL's survey paper [8] and summarize them here as follows:

Let

$$h : \mathbf{R}^m \to \mathbf{R}, \text{ with } h(x) \to \infty \text{ for } \|x\| \to \infty$$

be a convex and continous substitute for $\|.\|$ or $\|.\|^2$, bounded from below. We try to minimize $F(x) := h(f(x))$ for a function

$$f : \mathbf{R}^n \to \mathbf{R}^m, \;\; m \geq n, f \in C^1,$$

but in practice we minimize the **normalized linearization gain**

$$\psi(x) := F(x) - \inf_{\|s\| \leq 1, s \in \mathbf{R}^n} h(f(x) + \nabla f(x) s), \;\; x \in \mathbf{R}^n \tag{1}$$

and are content with stationary points, i.e. zeros of ψ. The analysis can be restricted to a domain $H \subset \mathbf{R}^m$ where

$$|h(f_1) - h(f_2)| \leq L \|f_1 - f_2\|, \;\; f_1, f_2 \in H \tag{2}$$

holds. Furthermore, the inequality

$$|F(x+s) - h(f(x) + \nabla f(x)s)|| \le L||s||\omega(s) \tag{3}$$

with the modulus $\omega(.)$ of continuity of ∇f is frequently used (see section 2 of [8]).

The algorithms we consider here will simply be defined as mappings

$$x \mapsto s(x), \quad \mathbf{R}^n \to \mathbf{R}^n$$

leading to the iteration

$$x_{k+1} := x_k + s(x_k). \tag{4}$$

A function $\Phi : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ will be needed to evaluate the progress of the algorithms. Standard choices will be (15) or (10), corresponding to POWELL's two convergence theorems in [8], but in principle the convergence theory does not depend on Φ.

2 General Convergence Theory

The following presentation of general convergence theorems combines POWELL's two approaches (for line–search– and trust–region methods) into one, extends them (e.g. to projection methods like DEUFLHARD's [3] and to linear convergence results) and includes the quantitative global convergence theorem of SCHABACK [9]. We keep the treatment fairly general, because we want to apply the results under different circumstances in a forthcoming paper.

The first axiom (see [8] and [9], for instance) concerns the descent of the error:

Axiom 2.1 *For all* $x \in \mathbf{R}^n$,

$$F(x + s(x)) \le (1-\gamma)F(x) + \gamma\Phi(x, s(x)), \gamma \in (0, 1). \tag{5}$$

The second axiom bounds $F(x) - \Phi(x, s(x))$ from below with respect to $\psi(x)$ (see [8]):

Axiom 2.2 *For all* $\delta > 0$ *there exists some* $\epsilon > 0$ *such that* $\psi(x) \ge \delta$ *implies*

$$F(x) - \Phi(x, s(x)) \ge \epsilon.$$

Theorem 2.1 *For the iteration* (4) *of a method satisfying Axioms 2.1 and 2.2,*

$$\psi(x_k) \to 0 \quad \text{for} \quad k \to \infty. \tag{6}$$

Proof: Whenever $\psi(x_k) \ge \delta, F(x_k) - F(x_{k+1}) \ge \epsilon$, but h and F are bounded from below. ■

This theorem combines Powell's theorems 1 and 2 from [8], while the next one incorporates a quantitative result of SCHABACK [9] based on the following

Axiom 2.3 *For all $x \in \mathbf{R}^n$*

$$F(x) - \Phi(x, s(x)) \geq \kappa g(\psi(x)) \tag{7}$$

for some $\kappa > 0$ and some nonnegative continous function $g : \mathbf{R} \to \mathbf{R}$ vanishing only at zero.

Theorem 2.2 *If Axioms 2.1 and 2.3 hold, then*

$$\sum_k g(\psi(x_k)) < \infty. \tag{8}$$

Proof: By summation of

$$F(x_k) - F(x_{k+1}) \geq \gamma(F(x_k) - \Phi(x_k, s(x_k))) \geq \gamma\kappa g(\psi(x_k)).\blacksquare$$

We shall apply this theorem in the cases $g(x) = x^2$ and x^4.

For linear convergence we use

Axiom 2.4

$$F(x) - \Phi(x, s(x)) \geq \kappa(F(x) - F(x^*)), \quad \kappa > 0 \tag{9}$$

for all x from a neighbourhood U of a critical point x^* of F.

This means that the algorithm achieves at least a fixed fraction of the total possible local error decrease.

Theorem 2.3 *Assume that $\{x_k\}$ has a limit point x^*. Then axioms 2.1 and 2.4 imply a linear convergence step*

$$F(x_{k+1}) - F(x^*) \leq (1 - \gamma\kappa)(F(x_k) - F(x^*))$$

whenever $x_k \in U$.

Proof: Obvious from (5) and (9). See also [9]. $\blacksquare$

3 A Class of Algorithms

We consider

$$\Phi(x, s) := h(f(x) + \nabla f(x)s) + \frac{1}{2}s^T B(x)s \tag{10}$$

with a symmetric $n \times n$ matrix $B(x)$ and

$$-\mu(x)\|s\|^2 \leq s^T B(x)s \leq M(x)\|s\|^2, \quad s \in \mathbf{R}^n \tag{11}$$

with certain functions $\mu(x)$, $M(x) : \mathbf{R} \to \mathbf{R}$, $M(x) \geq 0$. The algorithms are assumed to minimize $\Phi(x, s)$ with respect to s over some domain $S(x) \subset \mathbf{R}^n$ having the following properties:

1. $S(x) = P(x)\tilde{S}(x)$ for a linear projector $P(x)$ satisfying

$$\|\nabla f(x)(s - P(x)s)\| \le \sigma(x)\|s\|, \quad s \in \mathbf{R}^n$$

   for some function $\sigma(x) \ge 0$. This is done to cover methods like DEUFLHARD's (see e.g [3]). If $P(x)$ is not needed, $\sigma(x) = 0$, but it may improve the numerical behaviour of algorithms to project on "relevant" eigenspaces of $\nabla f^T(x)\nabla f(x)$ in case of rank loss of the Jacobian.

2. $\tilde{S}(x)$ contains a ball

$$K_{r(x)} := \{s \in \mathbf{R}^n \mid \|s\| \le r(x)\}, \quad r(x) \in (0, \infty]$$

   with $K_\infty = \mathbf{R}^n$.

3. $\tilde{S}(x)$ is contained in $K_{R(x)}, R(x) \in [r(x), \infty]$.

The solution $\hat{s}(x)$ of the minimization should exist (but must not necessarily be unique) and lie in $K_{\rho(x)}$:

$$\|\hat{s}(x)\| \le \rho(x), \quad 0 \le \rho(x) \le R(x). \tag{12}$$

We consider $R(x)$ as a controllable a-priori bound on the solution in the sense of trust-region methods, while $\rho(x)$ is an a-posteriori bound on the actual solution. For regularized line-search- and Levenberg-Marquardt methods, $R(x) = \infty$, but there are bounds $\rho(x)$ on $\|\hat{s}(x)\|$. For instance, Levenberg-Marquardt methods have

$$B(x) = M(x) \cdot Id, S(x) = \mathbf{R}^n, \mu(x) = -M(x), r(x) = R(x) = \infty,$$

and

$$\|\hat{s}(x)\|^2 \le \frac{2|F(x)|}{M(x)} \quad \text{for } M(x) > 0 \tag{13}$$

$$\|\hat{s}(x)\|^2 \le \frac{2|\psi(x)|}{M(x)} \quad \text{for } M(x) > 2|F(x)|. \tag{14}$$

To cover line-search methods, we finally take

$$s(x) := t(x)\hat{s}(x)$$

with a stepsize $t(x) \in (0, 1]$ to make the definition of our algorithm class complete.

This includes

- line search methods
- projection methods
- Levenberg-Marquardt methods
- trust region methods

together with a fairly free choice of second order terms $B(x)$. It will allow a theoretical comparison, because we shall give general conditions for convergence theorems like those in the preceding section. Note that in contrast to POWELL's approach we consider only "genuine" trust–region steps and regard adjustment of control parameters as an "inner" problem for each iteration.

The parameters $\mu(x), M(x), R(x), t(x)$, and $\sigma(x)$ are those the algorithm really uses at the step $x \mapsto x+s(x)$. However, any realistic implementation does not fix them beforehand, but rather updates them according to certain strategies. We assume that an implementation starts with a–priori parameters

$$\mu_0(x) \in \mathbf{R}, M_0(x) \geq 0, \sigma_0(x) \geq 0,$$

$$R_0(x) \geq r_0(x) > 0, t_0(x) \geq \hat{t} > 0,$$

and uses certain numerical tests to accept or modify these values. Modifications will always be by multiplication or division by a constant $\nu \in (0,1)$ until the test is passed. We leave the treatment of more sophisticated update formulae to the reader.

Furthermore, we restrict the control of parameters mainly to $R(x), t(x)$, and $\sigma(x)$. Controlling $\mu(x)$ or $M(x)$ is a common strategy only for Levenberg–Marquardt methods with positive definite $B(x)$ and implies a–posteriori bounds on $\tilde{s}(x)$ like (13) or (14). We consider such cases as indirect controls of $R(x)$ and give them a separate treatment.

4 Basic Lemmas

The convergence theory of section 2 can be applied for $\Phi(x,s)$ as in (10) or $\Phi(x,s)$ equal to

$$\varphi(x,s) = h(f(x) + \nabla f(x)s) \tag{15}$$

(see POWELL's two approaches in [8]). We use (15) just as an abbreviation and (10) for both minimization and convergence analysis.

We start with some vector $\bar{s}(x) \in \mathbf{R}^n$ giving a positive linearization gain

$$0 < \epsilon(x, \bar{s}(x)) \leq F(x) - \varphi(x, \bar{s}(x)). \tag{16}$$

The following lemmas can be applied for different choices of $\bar{s}(x)$:

Lemma 4.1 *If*

$$2\|\bar{s}(x)\|L\sigma(x) \leq \epsilon(x, \bar{s}(x)), \tag{17}$$

then

$$0 < \epsilon\left(x, P(x)\bar{s}(x)\right) := \frac{1}{2}\epsilon\left(x, \bar{s}(x)\right) \leq F(x) - \varphi\left(x, P(x)\bar{s}(x)\right). \tag{18}$$

Proof:

$$\begin{aligned} & F(x) - \varphi(x, P(x)\bar{s}(x)) \\ \geq\ & F(x) - h(f(x) + \nabla f(x)\bar{s}(x)) - L\|\nabla f(x)(\bar{s}(x) - P(x)\bar{s}(x))\| \\ \geq\ & \epsilon(x, \bar{s}(x)) - L\sigma(x)\|\bar{s}(x)\| \\ \geq\ & \tfrac{1}{2}\epsilon(x, \bar{s}(x)). \end{aligned}$$

■

Lemma 4.2 *If* (16) *and* (17) *hold, then for*

$$\tau(x) := \min\left(1, \frac{r(x)}{\|\bar{s}(x)\|}, \frac{\epsilon(x, \bar{s}(x))}{2M(x)\|\bar{s}(x)\|^2}\right) \tag{19}$$

we have

$$F(x) - \Phi(x, \tilde{s}(x)) \geq \frac{1}{4}\epsilon(x, \bar{s}(x))\tau(x). \tag{20}$$

If

$$(1 - t(x))\mu(x)\rho(x)^2 \leq t(x)\left(F(x) - \Phi(x, \tilde{s}(x))\right), \tag{21}$$

then

$$F(x) - \Phi(x, s(x)) \geq \frac{1}{2}t(x)\left(F(x) - \Phi(x, \tilde{s}(x))\right) \geq \frac{1}{8}\tau(x)t(x)\epsilon(x, \bar{s}(x)). \tag{22}$$

Proof: We omit the argument x, where possible, and get

$$\begin{aligned} \Phi(x, \tilde{s}) \leq\ & \Phi(x, \tau P\bar{s}) \\ =\ & h(f(x) + \nabla f(x)\tau P\bar{s}) + \tfrac{1}{2}\tau^2(P\bar{s})^T BP\bar{s} \\ \leq\ & F(x) - \tau\epsilon(x, P\bar{s}) + \tfrac{1}{2}\tau^2 M\|\bar{s}\|^2 \\ \leq\ & F(x) - \tfrac{1}{2}\tau\epsilon(x, \bar{s}) + \tfrac{1}{4}\tau\epsilon(x, \bar{s}), \end{aligned}$$

because $\tau P\bar{s}$ is admissible. Furthermore

$$\begin{aligned} \Phi(x, s) =\ & \Phi(x, t\tilde{s}) \\ =\ & h(f(x) + \nabla f(x)t\tilde{s}) + \tfrac{1}{2}t^2\tilde{s}^T B\tilde{s} \\ \leq\ & (1 - t)F(x) + t\Phi(x, \tilde{s}) - \tfrac{1}{2}t(1 - t)\tilde{s}^T B\tilde{s} \\ \leq\ & F(x) - t(F(x) - \Phi(x, \tilde{s})) + \tfrac{1}{2}(1 - t)\mu\rho^2 \\ \leq\ & F(x) - \tfrac{1}{2}t(F(x) - \Phi(x, \tilde{s})). \end{aligned}$$

■

Lemma 4.3 *The inequality*

$$Lt(x)\rho(x)\omega(t(x)\rho(x)) + \frac{1}{2}\mu(x)t^2(x)\rho^2(x) \leq (1 - \gamma)\left(F(x) - \Phi(x, s(x))\right) \tag{23}$$

implies (5).

Proof:

$$\begin{aligned} & (1 - \gamma)F(x) + \gamma\Phi(x, s) - h(f(x + s)) \\ \geq\ & (1 - \gamma)(F(x) - \Phi(x, s)) + \tfrac{1}{2}s^T Bs - L\|s\|\omega(\|s\|) \\ \geq\ & (1 - \gamma)(F(x) - \Phi(x, s)) - \tfrac{1}{2}\mu\rho^2 t^2 - Lt\rho\omega(t\rho) \\ \geq\ & 0. \end{aligned}$$

■

5 Global convergence

We now prove that any algorithm in our class will converge in the sense of Theorems 2.1 and 2.2, if parameters are controlled appropriately.

The analysis links $F(x) - \Phi(x, s(x))$ to $\psi(x)$ by using the vector $\hat{s}(x)$ maximizing $\psi(x)$, i.e.

$$\psi(x) = F(x) - h(f(x) + \nabla f(x)\hat{s}(x)), \quad \|\hat{s}(x)\| = 1.$$

We take $\bar{s}_1(x) = \hat{s}(x), \quad \bar{s}_2(x) = P(x)\bar{s}_1(x)$ and use Lemma 4.1 to get

$$\epsilon(x, \bar{s}_2(x)) \geq \frac{1}{2}\epsilon(x, \bar{s}_1(x)) = \frac{1}{2}\psi(x) \tag{24}$$

for bounds

$$\epsilon(x, \bar{s}_i(x)) \leq F(x) - h(f(x) + \nabla f(x)\bar{s}_i(x)),$$

if

$$2L\sigma(x) \leq \psi(x) \tag{25}$$

holds. Though L is unknown, a good choice of $\sigma_0(x)$ would be proportional to $\psi(x)$ for small $\psi(x)$. Using (24) as a test and multiplying $\sigma(x)$ by ν in case of failure will give a finite inner iteration that finally satisfies (24) and makes sure that the actually used parameter is bounded by

$$\sigma(x) \geq \min\left(\sigma_0(x), \frac{\nu\psi(x)}{2L}\right).$$

Now we apply Lemma 4.2 for $\bar{s}_2(x) = P(x)\hat{s}(x)$ and use $\|\bar{s}_2(x)\| \leq 1$ with

$$\tau(x) \geq \tau_1(x) := \min\left(1, r(x), \frac{\psi(x)}{2M(x)}\right) \tag{26}$$

to get

$$F(x) - \Phi(x, \tilde{s}(x)) \geq \frac{\psi(x)}{4}\tau_1(x) \tag{27}$$

(see POWELL's Lemma 5 in [8]). Then

$$(1 - t(x))\mu(x)\rho^2(x) \leq \frac{\psi(x)}{4}\tau_1(x)t(x) \tag{28}$$

will be sufficient for (22) to hold, and we get

$$F(x) - \Phi(x, s(x)) \geq \frac{\psi(x)}{8}\tau_1(x)t(x). \tag{29}$$

Methods with stepsize control that have no information about $\rho(x)$ will have problems with (28) in case of $\mu(x) > 0$ and $t(x) < 1$. This is the reason why genuine second-order methods (i.e.with possibly indefinite matrices $B(x)$) require a trust- region strategy or strong a-posteriori information about $\rho(x)$.

Now we apply Lemma 4.3 and use

$$L\rho(x)\omega(t(x)\rho(x)) + \frac{1}{2}\mu(x)t(x)\rho^2(x) \le \frac{1-\gamma}{8}\psi(x)\tau_1(x) \tag{30}$$

as a sufficient condition for (23) and (5), which can easily be satisfied by control of $t(x)$ or $R(x)$ after testing (5). Line-search methods will decrease $t(x)$ until

$$L\rho(x)\omega\left(\frac{t(x)}{\nu}\rho(x)\right) + \frac{1}{2}\mu(x)\frac{t(x)}{\nu}\rho^2(x) > \frac{1-\gamma}{8}\psi(x)\tau_1(x) \tag{31}$$

while trust-region methods guarantee $R(x) = R_0(x)$ or

$$L\frac{R(x)}{\nu}\omega\left(\frac{R^2(x)}{\nu}\right) + \frac{1}{2}\mu(x)\frac{R^2(x)}{\nu^2} > \frac{1-\gamma}{8}\psi(x)\tau_1(x). \tag{32}$$

The convergence of algorithms in the sense of Theorem 2.1 requires $\tau_1(x)$ and $t(x)$ to be bounded from below by positive expressions depending on $\psi(x)$ (see (29)). We state the following facts that can be read off the equations derived so far:

Theorem 5.1

1. *Using projections in any convergent algorithm is no problem provided that $\sigma(x)$ is controlled to satisfy (25).*
2. *Trust region methods (using $t(x) = 1, R(x) = r(x)$ and controlling $R(x)$ to ensure (5)) will converge if $M(x)$ and $\mu(x)$ are globally bounded from above by constants or singularities of type $\psi^{-\alpha}(x), \alpha > 0$.*
3. *Methods with absolutely no information on $\rho(x)$ (e.g. "pure" Gauss-Newton methods with $S(x) = \mathbf{R}^n$) cannot be proven to be convergent.*
4. *Methods with stepsize control will converge, if $\mu(x) = 0$ and if $M(x), R(x)$ are globally bounded.*

■

Levenberg–Marquardt type methods need a special treatment in this context. They assume $\mu(x) \le 0$ and therefore have no problems with (28) and (29). If they control $\mu(x) < 0$ by testing (5) and dividing $\mu(x)$ by ν in case of failure, they will guarantee (5) via (30) after finitely many iterations, because $\mu(x)$ tends to $-\infty$. If $\nu\mu(x)$ fails, still

$$-\frac{1}{2}\mu(x) < LK/\nu$$

in case of $\omega(t) \le Kt$ (e.g. if $f \in C^2$). Since the classical Levenberg–Marquardt method has $-\mu(x) = M(x)$, we can therefore assume a-posteriori that $M(x)$ and $\mu(x)$ are globally bounded away from infinity. Furthermore, the a-posteriori bounds (13) and (14) can be used as $\rho(x)$.

This gives a general version of the well-known result (see e.g. OSBORNE [7] or HÄUSSLER [4]):

Theorem 5.2 *Levenberg–Marquardt methods converge in the sense of Theorem 2.1.* ■

A revision of these arguments in view of (7) with $g(t) = t^2$ and Theorem 2.2 will reveal nearly the same facts for the quantitative convergence theory:

Theorem 5.3 *The methods considered so far are* **quantitatively** *globally convergent in the sense of*

$$\sum_k^\infty \psi^2(x_k) < \infty,$$

except that $M(x) = 0$ has to be assumed for line-search methods.

Proof: Examination of $F(x) - \Phi(x, s(x))$ reveals a $\mathcal{O}(\psi^2(x))$ behavior, except for the case of line–search methods, where $M(x) = 0$ is needed to avoid $\tau_1(x) \sim \psi(x)$ and $t(x) \sim \psi^2(x)$, which would otherwise lead to $\mathcal{O}(\psi^4(x))$. ■

The argument above shows that the case of $M(x) > 0$ for line–search methods can still be handled by using $g(t) = t^4$:

Theorem 5.4 *Well-controlled line-search methods with positive semidefinite second-order terms and globally bounded control parameters will give at least a convergence of*

$$\sum_k^\infty \psi^4(x_k) < \infty.$$

■

This illustrates how to apply our theory to certain examples from our class of algorithms. A general theory for **linear** convergence will be given in a forthcoming paper.

6 A specific method

We give an illustrative example of a "well–controlled" method whose convergence behavior can be analyzed easily, but which does not fall directly into the algorithm class described above. We (theoretically) minimize

$$\Phi(x, s) = \varphi(x, s) = h(f(x) + \nabla f(x)s)$$

subject to the "quadratic" constraint

$$K(x)\|s\|^2 \le (1 - \gamma)\left(F(x) - h(f(x) + \nabla f(x)s)\right). \qquad (33)$$

If $K(x)$ is not smaller than the constant K_0 appearing in

$$\frac{|h(f(x+s)) - h(f(x) + \nabla f(x)s)|}{\|s\|^2} \le K_0 \text{ for } x, s \in \mathbf{R}^n, s \neq 0 \qquad (34)$$

for $f \in C^2$, then (33) is a sufficient condition for (5), and Axiom 2.1 is satisfied.

A very simple control of $K(x)$ can therefore be implemented by testing (5) and multiplying the current (estimated or "last") value $k(x)$ for $K(x)$ by ν in case of failure. Estimates for $k(x)$ are cheaply available via quotients of the form arising in (34). Since results from optimization theory imply that the solution of the minimization subject to (33) is a solution of a generalized Levenberg–Marquardt regularization (in the sense of [9]), it will suffice to produce a Levenberg–Marquardt solution satisfying (33) with equality, using some constant $K \in [k(x), \nu k(x)]$ on the left–hand side. This solution can be obtained by a simple finite inner iteration of the associated Levenberg–Marquardt parameter, and for convergence analysis we can use the final value of K as $K(x)$ in (33).

This method can be viewed as a trust–region method using $t(x) = 1, \sigma(x) = \mu(x) = M(x) = 0$ and a dynamic control satisfying (5) directly; it gives

$$\rho^2(x) \leq \frac{1-\gamma}{K}(F(x) - \Phi(x, s(x))), \tag{35}$$

where K is a positive lower bound of $K(x)$.

It is now easy to prove that

$$F(x) - \Phi(x, s(x)) \geq \psi^2(x)(1-\gamma)K^{-1} \tag{36}$$

proceeding along the lines of the lemmas, and the global convergence theory implies

Theorem 6.1 *The method has quantitative global convergence in the sense of Theorem 2.2.* ■

The following result gives sufficient conditions for local linear convergence of this method:

Theorem 6.2 *The method has local linear convergence in parameter space, if around a limit point x^* the linearization gain of the optimal increment $x^* - x$ is at least quadratic and at least proportional to the total available error gain $F(x) - F(x^*)$:*

$$F(x) - \Phi(x, x^* - x) \geq c_0\|x - x^*\|^2, \quad c_0 > 0, \tag{37}$$

$$F(x) - \Phi(x, x^* - x) \geq c_1(F(x) - F(x^*)), \quad c_1 > 0. \tag{38}$$

Proof: It is easy to see from (37) that for a stepsize $t(x) \in (0, 1]$, bounded away from zero, the increment $s = t(x)(x^* - x)$ is admissible. Then (38) and

$$F(x) - \Phi(x, s(x)) \geq t(x)(F(x) - \Phi(x, x^* - x))$$

make sure that Axiom 2.4 is satisfied. Local linear convergence of $F(x_k)$ will then follow from Theorem 2.3. From (5) and (33) we deduce

$$K\|s(x_k)\|^2 \leq \frac{(1-\gamma)}{\gamma}(F(x_k) - F(x_{k+1}))$$

to get local linear convergence of $\|s(x_k)\| = \|x_k - x_{k+1}\|$ to zero. Standard arguments will then imply local linear convergence of $\{x_k\}$ to x^*. ■

7 Numerical Experience

The theoretical results derived so far (and those about local linear convergence to appear in a forthcoming paper) suggest that all well–controlled methods using no second–order terms should have about the same convergence behavior. This is supported by experience for a series of test runs with different methods and approximation problems (see Table 1 below).

All examples presented here have rather bad starting values (taken from the literature, if available), and in general the local convergence was surprisingly good. Reliable values for linear convergence factors were rare, and therefore a comparison by local linear convergence factors was impossible. Thus we restrict ourselves to simply report the number of function evaluations (gradient evaluations counted n–fold).

We used the same stopping criterion for all methods, namely $\psi(x) \leq 10^{-7}(1 + F(x))$, implemented by a shared and independent subroutine. Calculations were done in double precision on a VAX 11/780 under VMS–FORTRAN. The gradient of f was calculated by numerical differentiation by another shared subroutine.

The numerical results should be interpreted with caution. All methods spend most of their time far away from the optimum, and the variations in numerical effort depend to a large extent on pure chance, because the methods may happen to avoid or encounter regions where $\psi(x)$ is small, leading to poor progress in error reduction.

	L_2–Norm				L_∞–Norm	
Example	Deuflh	GN-Stp	Osb-LM	Quad.C	GN-Stp	Quad.C
BI	24	22	27	21	25	24
BPR	over	82	102	67	81	75
FR	106	44	153	304	130	124
JS	fail	54	21	49	52	43
KO	75	65	75	104	70	56
MR4	44	25	28	21	fail	21
MR5	45	26	28	24	27	24
MR6	52	84	124	138	98	134
OL1	48	54	306	266	51	58
OL2	121	108	605	341	139	over
W	fail	fail	403	81	136	131

Table 1: Function evaluations

Legend for Table 1:

Deuflh	:	DEUFLHARD's method NLSQN, [3]
GN-Stp	:	GAUSS–NEWTON method with stepsize control, [9]
Osb-LM	:	OSBORNE's implementation of the Levenberg–Marquardt algorithm, [7]
Quad.C	:	Algorithm with quadratic constraint (this paper)
fail	:	No convergence within 150 iterations
over	:	Overflow
BI	:	Example of BEN–ISRAEL, [4]
BPR	:	Example of BARRODALE, POWELL, and ROBERTS, [1] $\exp(x) - p_1(x)/q_3(x), n = 5$
FR	:	Example of FREUDENSTEIN and ROTH, [10], $n = 2$
JS	:	Example of JENNRICH and SAMPSON, [10], $n = 2$
KO	:	Example of KOWALIK and OSBORNE, [4], [5], [10]
MRx	:	Example of MEYER and ROTH No. x, [4], [6]
OL1	:	Example of OSBORNE, [10], $n = 5$
OL2	:	Example of OSBORNE, [10], $n = 12$
W	:	Example of WATSON, [10], $n = 12$

References

[1] Barrodale, I., Powell, M.J.D., and Roberts, F.D.K., The Differential Correction Algorithm for Rational l_∞-Approximation, SIAM J. Numer. Anal. 9, 493–504 (1972)

[2] Ben–Israel, A., A Newton–Raphson–Method for the solution of equations, J. Math. Anal. Appl. 15, 243–252 (1966)

[3] Deuflhard, P., Apostolescu, V., A study of the Gauss– Newton Algorithm for the Solution of Nonlinear Least Squares Problems, in: Frehse, J., Pallaschke, D., Trottenberg, U. (ed.): Special Topics of Applied Mathematics, 129–150, North Holland 1980

[4] Häußler, W.M., A Local Convergence Analysis for the Gauss–Newton and Levenberg–Morrison–Marquardt–Algorithms, Computing 31, 231–244 (1983)

[5] Kowalik, J., Osborne, M. R., Methods for unconstrained optimization problems. New York: American Elsevier 1968

[6] Meyer, R.R., and P.M. Roth, Modified damped least squares: an algorithm for non-linear estimation. J. Inst. Maths. Applics. 9, 218–233 (1972)

[7] Osborne, M.R., Nonlinear least squares: the Levenberg–Marquardt algorithm revisited, J. Austr. Math. Soc., Ser. B, 19, 343–357 (1972)

[8] Powell, M.J.D., General Algorithms for Discrete Nonlinear Approximation Calculations, in: Chui, Ch.K., Schumaker, L.L., and Ward, J.D. (eds): Approximation Theory IV, 187–218, Academic Press 1983

[9] Schaback, R., Convergence Analysis of the General Gauss–Newton Algorithm, Numer. Math. 46, 281–309 (1985)

[10] Watson, G.A., The Minimax Solution of an Overdetermined System of Non–linear Equations, J. Inst. Maths. Applics 23, 167–180 (1979)

Prof. Dr. R. Schaback
Institut für Numerische
und Angewandte Mathematik
der Universität Göttingen
Lotzestraße 16–18
D–3400–Göttingen